High-dimensional Big Data Exploration for Model Tuning and Evaluation

Hui Wan, Jonas Lukasczyk, David Rogers, Phil Rasch, Ross Maciejewski, and Hans Hagen

If you have a lot of simulations and model variables to analyze...

Our visualization experts have a tool that could help you navigate through the high-dimensional space – and save weeks/months of scripting time!

Contact: Hui Wan (Hui.Wan@pnnl.gov), Jonas Lukasczyk (jl@jluk.de)
Inconvenience of the AMWG Diagnostics Package

- Compares only 2 simulations at a time
- Cumbersome to show multiple fields (variables) at the same time

Wuyin has a tool to solve these problems, but…

- What if I have shorter simulations thus no “proper” climo files?
- What if I want to look at a field, a vertical level, or a cross-section that is not included in the current package?
The Wish

- An interactive, lightweight, easy-to-use visualization tool
- To help compare many fields simultaneously

- **Big data:** ~10 TB of model output

- **High-dimensional:**
  - Vertical levels
  - Days/months/seasons
  - Perturbed parameters
  - Different tuning simulations
  - Variables
Demo

• Using prototype at http://e.jluk.de/VisRaw.html
Our To-do List

• **Appearances and functionalities**
  – More intuitive layout; more color bars.
  – Additional interactions (e.g., selection of subdomain or cross-section by click-and-drag)
  – Zonal mean and other “usual views”
  – Parametric sensitivity analysis
  – Sharing and reproducibility

• **Practicalities**
  – A mechanism to host tens of TBs of data and run the java-based server
  – Parallel data processing and rendering

• **Any comments and suggestions are welcome!**